

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 10-207901

(43)Date of publication of application : 07.08.1998

(51)Int.Cl.

G06F 17/30
G06F 17/60

(21)Application number : 09-009797

(71)Applicant : NIPPON TELEGR & TELEPH CORP
<NTT>

(22)Date of filing : 22.01.1997

(72)Inventor : FUJIKAWA EIJI
YAMAMORI KAZUHIKO
KUBOSONO RIYUUJI

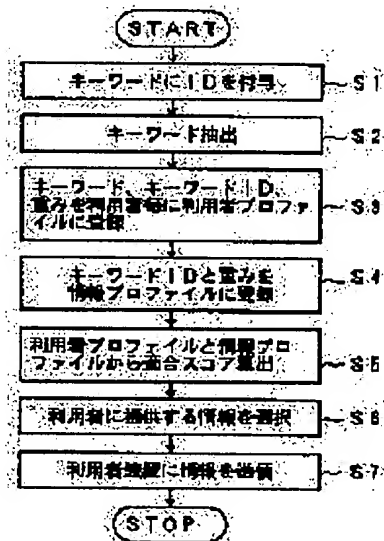
(54) METHOD AND SYSTEM FOR PROVIDING INFORMATION

(57)Abstract:

PROBLEM TO BE SOLVED: To shorten the time required for processing and to perform a higher processing by handling all of keywords of information, which is used to provide the profile of a user, as keyword IDs.

SOLUTION: An unequivocal number (keyword ID) is assigned to the character string of a keyword extracted from text information held in an equipment on the center side, and the keyword is extracted from text information (S1 and S2).

Keyword IDs corresponding to keywords whose weights are extracted as objects, which the user is interested in, and degrees of his interests in them are held in a user profile (S3). The weight of a keyword corresponding to a keyword ID is held in an information file, and the user profile and an information profile are used to calculate adaptation scores which indicate which degrees the information coincides with user's interest in by numerical values (S5), and information to be presented to a user device is selected in accordance with these scores (S6), and selected information is transmitted to the user device.



LEGAL STATUS

[Date of request for examination]

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平10-207901

(43) 公開日 平成10年(1998) 8月7日

(51) Int.Cl.⁴
G 0 6 F 17/30
17/60

識別記号

F I
G 0 6 F 15/403 3 4 0 A
15/21 Z
15/40 3 1 0 F

審査請求 未請求 請求項の数16 O L (全 15 頁)

(21) 出願番号 特願平9-9797

(22) 出願日 平成9年(1997) 1月22日

(71) 出願人 000004226

日本電信電話株式会社

東京都新宿区西新宿三丁目19番2号

(72) 発明者 藤川 英士

東京都新宿区西新宿三丁目19番2号 日本
電信電話株式会社内

(72) 発明者 山森 和彦

東京都新宿区西新宿三丁目19番2号 日本
電信電話株式会社内

(72) 発明者 窪田 竜二

東京都新宿区西新宿三丁目19番2号 日本
電信電話株式会社内

(74) 代理人 弁理士 伊東 忠彦

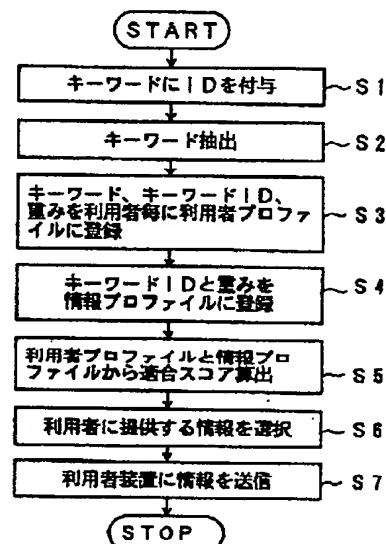
(54) 【発明の名称】 情報提供方法及びシステム

(57) 【要約】

【課題】 キーワードの文字列に一意の番号(キーワードID)を割当て、キーワードを全てキーワードIDで扱うことにより、適合スコアを算出する際に行わなければならないキーワードの検索の処理時間を減らすことが可能な情報提供方法及びシステムを提供する。

【解決手段】 センタ側装置に保持されているテキスト情報から抽出されたキーワードにキーワードIDを割当て、利用者の興味の高合として、重みを求め、抽出されたキーワードに対応するキーワードIDと共に利用者プロフィールに保持し、キーワードIDの重みを情報プロフィールに保持し、情報が利用者の興味にどの程度合致しているかを数値で示す適合スコアを算出し、適合スコアに従って利用者側装置に提供する情報を選択して送信する。

本発明の原理を説明するための図



【特許請求の範囲】

【請求項1】 情報が蓄積されたセンタ側装置中の個々の情報について、利用者の興味にどの程度合致しているかを定量的に算出し、その結果に基づいて情報を選択して、利用者側装置に提供する情報提供方法において、前記センタ側装置に保持されているテキスト情報から抽出されたキーワードの文字列に一意の番号（以下、キーワードIDと記す）を割当て、

前記テキスト情報からキーワードを抽出し、利用者が興味を持っている対象及び該対象に対する興味の度合としてのキーワードの重みを該キーワードに対応する前記キーワードIDと共に利用者プロフィールに保持し、

前記キーワードIDに対応するキーワードの重みを情報プロフィールに保持し、

前記利用者プロフィールと、前記情報プロフィールの内容を用いて、抽出されたキーワードに対応する前記情報が、前記利用者の興味にどの程度合致しているかを数値で示す適合スコアを算出し、

算出された前記適合スコアに従って前記利用者側装置に提供する情報を選択し、

選択された情報を前記利用者側装置に送信することを特徴とする情報提供方法。

【請求項2】 前記情報内のキーワードの文字列に前記キーワードIDを割当て、第1のテーブルに保持する請求項1記載の情報提供方法。

【請求項3】 前記選択された情報を前記利用者側装置に提供する際に、

前記適合スコアに従って利用者に提供する情報を前記利用者側装置に表示可能な情報に編集する請求項1記載の情報提供方法。

【請求項4】 前記選択された情報を前記利用者側装置に提供後、前記利用者側装置において前記利用者が行った操作の履歴を記憶しておく、

前記操作の履歴を用いて、提供した情報が利用者の興味にどれほど適していたかを適合度として算出して、保持し、

個々の情報についての適合度を参照して、前記利用者プロフィールにキーワードIDの追加または、削除を行う、または、該利用者プロフィール中のキーワードの重みを更新する請求項1記載の情報提供方法。

【請求項5】 情報が蓄積されたセンタ側装置中の個々の情報について、利用者の興味にどの程度合致しているかを定量的に算出し、その結果に基づいて情報を選択して、利用者側装置に提供する情報提供方法において、前記センタ側装置に保持されているテキスト情報のキーワードの文字列に一意の番号（以下、キーワードIDと記す）を割当て、

前記センタ側装置が保持する前記テキスト情報からキーワードを抽出し、

前記センタ側装置に蓄積された個々のテキスト情報におけるキーワードの出現回数を数え、該キーワードに対応するキーワードIDと共に情報プロフィールに保持しておく、

利用者が興味を持っている対象及び該対象に対する興味の度合として、キーワードに対応する前記キーワードIDを該キーワードの重みと共に利用者プロフィールに保持し、

前記センタ側装置に蓄積された全前記テキスト情報または、量的に十分なテキスト情報におけるキーワードの出現する割合を示す出現頻度を求め、キーワードIDと共に保持しておく、

前記情報プロフィールの前記出現回数と前記出現頻度から、該キーワードの重みを算出し、

前記利用者プロフィールの内容と、前記抽出されたキーワードに対する重みに基づいて適合スコアを算出し、

前記適合スコアに従って前記利用者側装置に提供する情報を選択し、

選択された情報を前記利用者側装置に送信することを特徴とする情報提供方法。

【請求項6】 前記情報内のキーワードの文字列に前記キーワードIDを割当て、第1のテーブルに保持する請求項5記載の情報提供方法。

【請求項7】 前記選択された情報を前記利用者側装置に提供する際に、

前記適合スコアに従って利用者に提供する情報を前記利用者側装置に表示可能な情報に編集する請求項5記載の情報提供方法。

【請求項8】 前記選択された情報を前記利用者側装置に提供後、前記利用者側装置において前記利用者が行った操作の履歴を記憶しておく、

前記操作の履歴を用いて、提供した情報が利用者の興味にどれほど適していたかを適合度として算出して、保持し、

個々の情報についての適合度を参照して、前記利用者プロフィールにキーワードIDの追加または、削除を行う、または、該利用者プロフィール中のキーワードの重みを更新する請求項5記載の情報提供方法。

【請求項9】 情報が蓄積された個々の情報について、利用者の興味にどの程度合致しているかを定量的に算出し、その結果に基づいて情報を選択して、利用者側装置に提供するセンタ側装置と、該センタ側装置から該情報を取得する利用者側装置からなる情報提供システムであって、

前記センタ側装置は、

利用者に提供するための情報を格納する情報蓄積手段と、

前記情報蓄積手段からキーワードの抽出及び抽出された該キーワードに重みを付与し、該キーワードに対応するキーワードIDを情報プロフィールに登録する情報プロ

ファイル生成手段と、

利用者毎にキーワードに対応する情報への関心度として、該キーワードに対応するキーワードIDと共に重みを保持する利用者情報保持手段と、

前記情報プロフィールと前記利用者情報保持手段の内容に基づいて、抽出されたキーワードに対応する情報が利用者の興味にどの程度合致しているかの度合である適合スコアを算出する適合スコア算出手段と、

前記適合スコア算出手段により求められた前記適合スコアに応じて、前記利用者側装置に送信する情報を選択する情報選択手段と、

前記情報選択手段で選択され情報を前記利用者側装置に送信する送信手段を有することを特徴とする情報提供システム。

【請求項10】 キーワードの文字列に一意に対応するキーワードIDを付与し、保持するキーワードIDテーブルを更に有し、

情報プロフィール生成手段は、前記キーワードテーブルを参照して、前記キーワードに対応するキーワードIDを取得する手段を含む請求項9記載の情報提供システム。

【請求項11】 前記利用者側装置は、

前記センタ側装置から送信された情報について操作した履歴情報を保持する操作履歴保持手段と、

前記操作履歴保持手段に保持されている前記履歴情報を前記センタ側装置に所定の周期、または、任意の周期で転送する履歴情報転送手段とを有し、

前記センタ側装置は、

前記利用者側装置から取得した前記履歴情報に基づいて、前記送信手段により送信した情報が利用者の興味にどれほど適していたかを示す適合度を算出する適合度算出手段と、

前記適合度算出手段により求められた前記適合度に基づいて、前記利用者情報保持手段の内容を更新する利用者情報更新手段を更に有する請求項9記載の情報提供システム。

【請求項12】 前記利用者情報更新手段は、

前記適合度に基づいて、前記利用者情報保持手段のキーワードIDを追加・削除、または、キーワードIDの重みを変更する手段を含む請求項11記載の情報提供システム。

【請求項13】 情報が蓄積された個々の情報について、利用者の興味にどの程度合致しているかを定量的に算出し、その結果に基づいて情報を選択して、利用者側装置に提供するセンタ側装置と、該センタ側装置から該情報を取得する利用者側装置からなる情報提供システムであって、

前記センタ側装置は、

利用者に提供するための情報を格納する情報蓄積手段と、

前記情報蓄積手段からキーワードを抽出し、抽出された該キーワードの該情報蓄積手段中の個々の情報における出現回数をカウントし、前記キーワードに対応するキーワードIDと共に保持する情報プロフィール生成手段と、

利用者毎にキーワードに対応する情報への関心度として、該キーワードに対応するキーワードIDと共に重みを保持する利用者情報保持手段と、

前記情報蓄積手段に蓄積されている全情報または、量的に十分な情報における任意のキーワードの出現頻度を、該キーワードに対応するキーワードIDと共に保持するキーワード出現頻度保持手段と、

抽出されたキーワードについて、該キーワードに対応する情報における出現回数と出現頻度を用いて重みを算出し、該重みと前記利用者情報保持手段の内容に基づいて、抽出されたキーワードに対応する情報が利用者の興味にどの程度合致しているかの度合である適合スコアを算出する適合スコア算出手段と、

前記適合スコア算出手段により求められた前記適合スコアに応じて、前記利用者側装置に送信する情報を選択する情報選択手段と、

前記情報選択手段で選択され情報を前記利用者側装置に送信する送信手段を有することを特徴とする情報提供システム。

【請求項14】 キーワードの文字列に一意に対応するキーワードIDを付与し、保持するキーワードIDテーブルを更に有し、

情報プロフィール生成手段は、前記キーワードテーブルを参照して、前記キーワードに対応するキーワードIDを取得する手段を含む請求項13記載の情報提供システム。

【請求項15】 前記利用者側装置は、

前記センタ側装置から送信された情報について操作した履歴情報を保持する操作履歴保持手段と、

前記操作履歴保持手段に保持されている前記履歴情報を前記センタ側装置に所定の周期、または、任意の周期で転送する履歴情報転送手段とを有し、

前記センタ側装置は、

前記利用者側装置から取得した前記履歴情報に基づいて、前記送信手段により送信した情報が利用者の興味にどれほど適していたかを示す適合度を算出する適合度算出手段と、

前記適合度算出手段により求められた前記適合度に基づいて、前記利用者情報保持手段の内容を更新する利用者情報更新手段を更に有する請求項13記載の情報提供システム。

【請求項16】 前記利用者情報更新手段は、

前記適合度に基づいて、前記利用者情報保持手段のキーワードIDを追加・削除、または、キーワードIDの重みを変更する手段を含む請求項15記載の情報提供システム。

テム。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、情報提供方法及びシステムに係り、特に、情報が蓄積されたセンタ側装置の個々の情報について、利用者の興味にどの程度合致しているかを定量的に算出し、その結果に基づいて情報を選択して、利用者側の装置に提供する情報提供方法及びシステムに関する。

【0002】

【従来の技術】従来において、情報が蓄積されたセンタ側装置中の個々の情報について、利用者の興味にどの程度合致しているかを定量的に算出し、その結果に基づいて情報を選択して利用者側装置に提供する情報提供方法及び装置においては、利用者の興味の対象として、予めキーワードや情報が属するジャンル等のインデックスを登録しておき、登録されたキーワードあるいは、インデックスによってセンタ側に蓄積された情報を検索し、検索の結果得られた情報を利用者の興味に合致した情報として提供する。

【0003】しかし、上記の方法では、利用者が欲しい情報を検索するためのキーワードを全て事前に登録しておくことは難しく、利用者が欲しい情報を網羅的に提供することができないという問題が残る。そこで、情報を提供したときの利用者の行動（検索するときに用いるキーワードや、情報の閲覧時間等）から利用者がその情報に対してどれだけ興味を持っていたかを推測し、その度合に従ってその情報から抽出されたキーワードを自動的に登録する方法がある。

【0004】しかし、AND検索やOR検索のように登録されているキーワードのいくつかを組み合わせて検索することができないという問題がある。これに対して、登録されているキーワードを組み合わせて検索を行い、情報を提供するため、利用者が興味を持っている対象に対する興味の度合として、キーワードとキーワードの重みを保持し、情報からキーワードを抽出し、その情報に対するそのキーワードの重要度としてキーワードとキーワードの重みを保持し、これらのキーワードの重みに関する計算を行うことによって、その情報が利用者の興味とどれだけ合致しているかを数値で示す適合スコアを算出する方法がある。

【0005】また、従来の方法では、利用者の興味の対象としてキーワードを登録する際に、キーワードを文字列データとして保存している。

【0006】

【発明が解決しようとする課題】しかしながら、上記従来の情報提供方法では、登録されている全てのキーワードを用いて、全ての情報から抽出されたキーワードを検索する処理が必要となり、処理に時間を要するという問題がある。例えば、利用者が興味を持っている対象とし

て、100個のキーワードとその重みを保持し、1万件の蓄積された情報についてそれぞれ平均20個のキーワードが抽出された場合、利用者の興味と情報との合致を調べるためには2千万回の検索を繰り返す必要があるが、従来のように、キーワードを文字列データとして扱う方法では、キーワード毎にデータの大きさが異なるため、画一的に扱うことができず、処理に時間を要するという問題がある。

【0007】本発明は、上記の点に鑑みなされたもので、キーワードの文字列に一意の番号（キーワードID）を割当て、キーワードを全てキーワードIDで扱うことにより、適合スコアを算出する際に行わなければならないキーワードの検索の処理時間を減らすことが可能な情報提供方法及びシステムを提供することを目的とする。

【0008】

【課題を解決するための手段】図1は、本発明の原理を説明するための図である。本発明は、情報が蓄積されたセンタ側装置中の個々の情報について、利用者の興味にどの程度合致しているかを定量的に算出し、その結果に基づいて情報を選択して、利用者側装置に提供する情報提供方法において、センタ側装置に保持されているテキスト情報から抽出されたキーワードの文字列に一意の番号（以下、キーワードIDと記す）を割当て（ステップ1）、テキスト情報からキーワードを抽出し（ステップ2）、利用者が興味を持っている対象及び該対象に対する興味の度合であるキーワードの重みを抽出されたキーワードに対応するキーワードIDと共に利用者プロフィールに保持し（ステップ3）、キーワードIDに対応するキーワードの重みを情報プロフィールに保持し（ステップ4）、利用者プロフィールと、情報プロフィールを用いて、情報が利用者の興味にどの程度合致しているかを数値で示す適合スコアを算出し（ステップ5）、適合スコアに従って利用者側装置に提供する情報を選択し（ステップ6）、選択された情報を利用者側装置に送信する（ステップ7）。

【0009】さらに、本発明は、情報が蓄積されたセンタ側装置中の個々の情報について、利用者の興味にどの程度合致しているかを定量的に算出し、その結果に基づいて情報を選択して、利用者側装置に提供する情報提供方法において、センタ側装置に保持されているテキスト情報のキーワードの文字列に一意の番号（以下、キーワードIDと記す）を割当て、センタ側装置が保持するテキスト情報からキーワードを抽出し、利用者が興味を持っている対象及び該対象に対する興味の度合であるキーワードの重みと、抽出されたキーワードに対応するキーワードIDと共に利用者プロフィールに保持し、センタ側装置に蓄積された個々のテキスト情報におけるキーワードの出現回数を数え、該キーワードに対応するキーワードIDと共に情報プロフィールに保持しておき、セン

タ側装置に蓄積された全テキスト情報または、量的に十分なテキスト情報におけるキーワードの出現する割合を示す出現頻度を求め、キーワードIDと共に保持しておき、抽出されたキーワードの出現回数と出現頻度から、該キーワードの重みを算出し、利用者プロフィールの内容と、抽出されたキーワードに対する重みに基づいて適合スコアを算出し、適合スコアに従って利用者側装置に提供する情報を選択し、選択された情報を利用者側装置に送信する。

【0010】また、本発明は、情報内のキーワードの文字列にキーワードIDを割当て、第1のテーブルに保持する。また、本発明は、選択された情報を利用者側装置に提供する際に、適合スコアに従って利用者に提供する情報を利用者側装置に表示可能な情報に編集する。

【0011】また、本発明は、選択された情報を利用者側装置に提供後、利用者側装置において利用者が行った操作の履歴を記憶しておき、操作の履歴を用いて、提供した情報が利用者の興味にどの程度適していたかを適合度として算出して、保持し、個々の情報についての適合度を参照して、利用者プロフィールにキーワードIDの追加または、削除を行う、または、該利用者プロフィール中のキーワードの重みを更新する。

【0012】図2は、本発明の原理構成図である。本発明は、情報が蓄積された個々の情報について、利用者の興味にどの程度合致しているかを定量的に算出し、その結果に基づいて情報を選択して、利用者側装置に提供するセンタ側装置10と、該センタ側装置10から該情報を取得する利用者側装置20からなる情報提供システムであって、センタ側装置10は、利用者に提供するための情報を格納する情報蓄積手段11と、情報蓄積手段11からキーワードの抽出及び抽出された該キーワードに重みを付与し、該キーワードに対応するキーワードIDを情報プロフィール14に登録する情報プロフィール生成手段12と、利用者毎にキーワードに対応する情報への関心度として、該キーワードに対応するキーワードIDと共に重みを保持する利用者情報保持手段15と、情報プロフィール14と利用者情報保持手段15の内容に基づいて、抽出されたキーワードに対応する情報が利用者の興味にどの程度合致しているかの度合である適合スコアを算出する適合スコア算出手段16と、適合スコア算出手段16により求められた適合スコアに応じて、利用者側装置に送信する情報を選択する情報選択手段17と、情報選択手段17で選択された情報を利用者側装置20に送信する送信手段40とを有する。

【0013】また、本発明は、キーワードの文字列に一意に対応するキーワードIDを付与し、保持するキーワードIDテーブル13を更に有し、情報プロフィール生成手段12は、キーワードテーブル13を参照して、キーワードに対応するキーワードIDを取得する手段を含む。

【0014】また、上記の利用者側装置20は、センタ側装置10から送信された情報について操作した履歴情報を保持する操作履歴保持手段23と、操作履歴保持手段23に保持されている履歴情報をセンタ側装置10に所定の周期、または、任意の周期で転送する履歴情報転送手段24とを有し、センタ側装置10は、利用者側装置から取得した履歴情報に基づいて、送信手段により送信した情報が利用者の興味にどの程度適していたかを示す適合度を算出する適合度算出手段18と、適合度算出手段18により求められた適合度に基づいて、利用者情報保持手段15の内容を更新する利用者情報更新手段19を更に有する。

【0015】上記の利用者情報更新手段19は、適合度に基づいて、利用者情報保持手段15のキーワードIDを追加・削除、または、キーワードIDの重みを変更する手段を含む。さらに、本発明は、情報が蓄積された個々の情報について、利用者の興味にどの程度合致しているかを定量的に算出し、その結果に基づいて情報を選択して、利用者側装置に提供するセンタ側装置と、該センタ側装置から該情報を取得する利用者側装置からなる情報提供システムであって、センタ側装置は、利用者に提供するための情報を格納する情報蓄積手段と、情報蓄積手段からキーワードを抽出し、抽出された該キーワードの該情報蓄積手段中の個々の情報における出現回数をカウントし、キーワードに対応するキーワードIDと共に保持する情報プロフィール生成手段と、利用者毎にキーワードに対応する情報への関心度として、該キーワードに対応するキーワードIDと共に重みを保持する利用者情報保持手段と、情報蓄積手段に蓄積されている全情報または、量的に十分な情報における任意のキーワードの出現頻度を、該キーワードに対応するキーワードIDと共に保持するキーワード出現頻度保持手段と、抽出されたキーワードについて、該キーワードに対応する情報における出現回数と出現頻度を用いて重みを算出し、該重みと利用者情報保持手段の内容に基づいて、抽出されたキーワードに対応する情報が利用者の興味にどの程度合致しているかの度合である適合スコアを算出する適合スコア算出手段と、適合スコア算出手段により求められた適合スコアに応じて、利用者側装置に送信する情報を選択する情報選択手段と、情報選択手段で選択され情報を利用者側装置に送信する送信手段を有する。

【0016】上記のように、本発明によれば、テキスト情報に対して、任意のキーワード抽出方式と任意のキーワード重み付与方式を用いつつ、キーワードにキーワードIDを割当て、利用者のプロフィール中のキーワードと蓄積された情報のキーワードを全てキーワードIDとして扱うことによって、適合スコアを算出する際に、利用者プロフィール中のキーワードから蓄積されているすべての情報を検索し、その情報でそのキーワードの重みを参照する処理に要する時間を短縮することが可能とな

る。

【0017】また、本発明によれば、テキスト情報に対して、キーワードの出現頻度を用いたキーワードの重みを付与する方式を用いつつ、キーワードをキーワードIDに変換し、利用者のプロフィールを提供するための情報のキーワードを全てキーワードIDとして取り扱うことによって、適合スコアを算出する際に、利用者プロフィール中のキーワードIDから蓄積されている全ての情報を検索し、その情報でそのキーワードの重みを参照する処理に要する時間を短縮することが可能となる。

【0018】

【発明の実施の形態】最初に第1のセンタ側装置の構成について説明する。図3は、本発明のセンタ側装置の構成図(その1)である。同図に示すセンタ側装置10は、情報蓄積部11、情報解析部12、キーワードIDテーブル13、情報プロフィール14、利用者プロフィール15、適合スコア算出部16、提供情報作成部17、適合度算出部18、利用者プロフィール修正部19から構成される。

【0019】情報蓄積部11は、利用者に提供するための個々のテキスト情報を格納する。情報解析部12は、キーワードの出現回数、出現位置、文字数等を参照して、キーワードを抽出し、キーワードの重みを算出する。なお、キーワードの重みを算出する方法は、公知の方法を用いるものとし、本発明ではその方法を規定しない。

【0020】キーワードIDテーブル13は、キーワードと当該キーワードに一意に付与されるキーワードIDが格納される。図4は、本発明のキーワードIDテーブルのデータ構造の例である。同図に示すキーワードIDテーブル13は、キーワードID131と実際のキーワードの文字列データ132が格納されている。

【0021】情報プロフィール14は、情報蓄積部11に蓄積された個々の情報について、情報解析部12によって抽出されたキーワードを、キーワードIDテーブル13を参照して、キーワードIDに変換し、キーワードの重みと共に保持する。図5は、本発明の情報プロフィールのデータ構造の例である。

【0022】同図に示す情報プロフィール14は、キーワードID141と当該キーワードIDに対応するキーワードの重み142から構成される。利用者プロフィール15は、利用者が興味を持っている対象及びその対象に対する興味の度合としてキーワードIDとその重みを保持する。

【0023】図6は、本発明の利用者プロフィールのデータ構造の例である。同図に示す利用者プロフィール15は、キーワードID151と、キーワードの重み152から構成される。この例では、データの大きさを統一して検索速度を速くするため、キーワードID151に固定長の16進数値を用いている。

【0024】適合スコア算出部16は、抽出されたキーワードに対する情報が利用者の興味にどの程度合致しているかを示す数値である適合スコアを情報プロフィール14と利用者プロフィール15から算出する。適合スコアの詳細な算出方法については後述する。

【0025】提供情報作成部17は、適合スコアに従って利用者に提供する情報を選択し、利用者側装置に表示するための選択した情報を加工及び編集する。適合度算出部18は、利用者側から送信された利用履歴テーブルの情報に従って、提供した個々の情報についての適合度を算出する。適合度の詳細な算出方法については後述する。

【0026】利用者プロフィール修正部19は、適合度算出部18により求められた個々の情報についての適合度に基づいて利用者プロフィール15の内容を修正する。次に、利用者側装置の構成について説明する。図7は、本発明の利用者側装置の構成図である。

【0027】同図に示す利用者側装置20は、情報表示部21、操作部22及び履歴情報保持部23から構成される。情報表示部21は、センタ側装置10から送られてきた情報を表示する。操作部22は、表示した情報に対する利用者の保存、破棄、画面スクロール等の操作を受け付ける。

【0028】履歴情報保持部22は、センタ側装置10から提供された情報の閲覧時間、情報の保存、破棄、画面のスクロール等の利用者の操作を観察し、それらの操作履歴をテーブル(利用履歴テーブル)として保存する。利用履歴テーブルには、例えば、提供された情報とその情報に対する閲覧時間とを設定する。

【0029】次に、センタ側装置10の動作を説明する。図8は、本発明の一連の動作の概要を示すフローチャートである。

ステップ100) 最初に、センタ側装置10において、情報解析部12が情報蓄積部11に蓄積されている情報からキーワードの出現回数、出現位置、文字数等を参照してキーワードを抽出し、当該キーワードに対応するキーワードIDをキーワードIDテーブル13から取得し、当該キーワードの重みと共に情報プロフィール14に設定する。

【0030】ステップ110) 適合スコア算出部16は、利用者毎に設定されているキーワードIDとそのキーワードIDに対応するキーワードに関する重みを利用者プロフィール15から取得し、さらに、情報プロフィール14の情報から、抽出されたキーワードに対応する情報の適合スコアを算出する。

【0031】ステップ120) 提供情報作成部17は、適合スコアに従って、利用者側装置20に送信する情報を選択し、利用者側装置20で表示可能な情報に編集し、編集された情報を利用者側装置20に送信する。

ステップ130) 利用者側装置20では、センタ側装

置10から取得した情報を表示し、当該表示情報に対する操作を行い、その操作状況を操作履歴テーブルに設定する。所定の周期または、任意のタイミングで操作履歴テーブルの情報をセンタ側装置10に転送する。

【0032】ステップ140) センタ側装置10は操作履歴テーブルの情報を受け取ると、当該情報に基づいて適合度を算出し、利用者プロフィール修正部19において、当該適合度に基づいて利用者プロフィール15を更新する。次に、上記のステップ100における情報プロフィール14を作成する処理について説明する。

【0033】図9は、本発明の第1のセンタ側装置における情報プロフィール作成処理のフローチャートである。

ステップ101) センタ側装置10に接続される入力部(図示せず)から利用者側装置20に提供するためのテキスト情報の入力を受け付け、情報蓄積部11に格納する。

【0034】ステップ102) 情報解析部12は、入力され、情報蓄積部11に蓄積されているテキスト情報から、既知の方法によるキーワード抽出処理及び、キーワード重み算出処理を行う。

ステップ103) キーワードIDテーブル13に抽出されたキーワードが登録されているかを判定し、登録されている場合にはステップ105に移行し、登録されていない場合にはステップ104に移行する。

【0035】ステップ104) キーワードIDテーブル13に使用されていないキーワードIDを抽出されたキーワードに対して割当て、当該キーワードと共にキーワードIDをキーワードIDテーブル13に登録する。

ステップ105) 情報解析部12は、抽出したキーワードをキーワードIDテーブルを参照して、キーワードIDに変換し、キーワードIDとステップ102において求められたキーワードに対する重みを情報に対応する情報プロフィール14に登録する。

【0036】次に、上記のステップ110におけるセンタ側装置10の適合スコア算出処理について説明する。図10は、本発明の適合スコア算出処理のフローチャートである。

ステップ111) 適合スコア算出部16は、各利用者について、利用者に対する適合スコアをまだ算出していない情報の情報プロフィール14の内容を取得する。

【0037】ステップ112) 情報プロフィール14と利用者プロフィール15によって適合スコアを算出する。適合スコアの算出は、以下のように行う。キーワードIDテーブル13に登録されているキーワード(N個とする)を一次独立なベクトルとするN次元ベクトル空間について考える。

【0038】利用者プロフィール15に登録されているキーワードをこのベクトル空間の部分空間と考え、キーワードの重み u_i ($i=1, 2, \dots, N$)を各要素の大

きさと考えて利用者プロフィール15のベクトル P_u を以下のように定義する。

$$P_u = (u_1, u_2, u_3, \dots, u_N)^T$$

但し、登録されていないキーワードIDについての重みは0とする。また、Tは、転置を表す。

【0039】同様に、情報プロフィール P_I についても定義する。但し、情報中のキーワードの重みを v_j ($j=1, 2, \dots, N$)とする。

$$P_I = (v_1, v_2, v_3, \dots, v_N)^T$$

但し、登録されていないキーワードIDについての重みは0とする。

【0040】このとき、適合スコアSを以下のように定義する。

$$S = P_u \cdot P_I$$

但し、式中の記号 \cdot は、内積演算を表す。

ステップ113) 提供情報作成部17は、適合スコアの順に情報のタイトルあるいは、要約情報等をリストにし、適合スコアを付与するなど、利用者に提供する情報を作成する。

【0041】ステップ114) 作成した情報を利用者側装置20に送信する。次に、上記のステップ140における利用者プロフィール修正処理の動作を説明する。図11は、本発明の利用者プロセスファイルの修正処理のフローチャートである。

【0042】ステップ141) センタ側装置10において、利用者側装置20から利用者が情報を閲覧した時間等を記した操作履歴テーブルの内容を取得する。

ステップ142) 適合度算出部18は、取得した操作履歴テーブルの履歴情報に基づいて利用者側装置20に提供したそれぞれの情報についての適合度を計算する。

【0043】適合度をaとすると、例えば、提供した情報に対して利用者が可もなく不可もなくという反応をみせたときの閲覧時間を、

$$a=0$$

として基準時間とし、

$a>0$: 基準時間より多くの時間を情報の閲覧に費やした

$a<0$: 基準時間より少ない時間を情報の閲覧に費やした

というように定義し、実験による経験的に求めた値としてaを決定する。

【0044】ステップ143) 利用者プロフィール修正部19は、ステップ142で求められた適合度を用いて、利用者プロフィール15を修正する。利用者プロフィール15の修正は、修正後の利用者プロフィール P_u' とすると、

$$P_u' = P_u + aP_I$$

と計算する。

【0045】次に、キーワード出現頻度を用いて適合スコアを算出する機能を有するセンタ側装置10について

説明する。図12は、本発明のセンタ側装置の構成図（その2）である。同図において、図3と同一構成部分には同一符号を付す。

【0046】同図に示すセンタ側装置10は、情報蓄積部11、情報解析部12、キーワードIDテーブル13、情報プロフィール14、利用者プロフィール15、適合スコア算出部16、提供情報作成部17、適合度算出部18、利用者プロフィール修正部19及びキーワード出現頻度記録部31から構成される。

【0047】情報蓄積部11は、利用者に提供するための個々のテキスト情報を格納する。情報解析部12は、キーワードの出現回数、出現位置、文字数等を参照して、キーワードを抽出し、その出現回数をカウントする。キーワードIDテーブル13は、キーワードと当該キーワードに一意に付与されるキーワードIDが格納される。

【0048】キーワード出現頻度記録部31は、センタ側装置10の情報蓄積部11に蓄積された全テキスト情報あるいは、量的に十分なテキスト情報におけるキーワードが出現する割合を示す出現頻度を保持する。図13は、本発明のキーワードIDテーブルとキーワード出現頻度記録を組み合わせたデータ構造の例である。

【0049】同図に示す構造は、キーワードID32、キーワード33、及び当該キーワード33に関する情報蓄積部11の全テキスト情報または、量的に十分なテキスト情報におけるキーワード出現頻度34から構成される。情報プロフィール14は、情報蓄積部11に蓄積された個々の情報について、情報解析部12によって抽出されたキーワードを、キーワードIDテーブル13を参照して、キーワードIDに変換し、キーワードの重みと共に保持する。

【0050】図14は、本発明の第2の情報プロフィールのデータ構造の例である。同図に示す情報プロフィール14は、キーワードID143と当該キーワードIDに対応するキーワードの出現回数144から構成される。この例では、キーワードIDに16進数を用いている。

【0051】利用者プロフィール15は、利用者が興味を持っている対象及びその対象に対する興味の度合としてキーワードIDとその重みを保持する。適合スコア算出部16は、抽出されたキーワードに対する情報が利用者の興味にどの程度合致しているかを示す数値である適合スコアを情報プロフィール14のキーワードの出現回数144とキーワード出現頻度記録部31からその情報に対する重要度として重みを算出し、利用者プロフィール15を用いて適合スコアを算出する。

【0052】提供情報作成部17は、適合スコアに従って利用者に提供する情報を選択し、利用者側装置に表示するための選択した情報を加工及び編集する。適合度算出部18は、利用者側から送信された利用履歴テーブル

に従って、提供した個々の情報についての適合度を算出する。

【0053】利用者プロフィール修正部19は、適合度算出部18により求められた個々の情報についての適合度に基づいて利用者プロフィール15を修正する。次に、図8におけるステップ100の情報プロフィール14の作成処理について説明する。

【0054】図15は、本発明の第2のセンタ側装置における情報プロフィールの作成処理のフローチャートである。

ステップ301) センタ側装置10に接続される入力部（図示せず）からの情報蓄積部11への利用者に提供するためのテキスト情報の入力を受け付ける。

【0055】ステップ302) 情報解析部12は、入力されたテキスト情報からキーワードの抽出を行い、キーワードの出現回数をカウントする。

ステップ303) 抽出したキーワードがキーワードIDテーブル13に登録されているかを調べ、登録されていない場合には、ステップ304に移行し、登録されている場合には、ステップ306に移行する。

【0056】ステップ304) キーワードIDテーブル13に使用されていないキーワードIDを抽出されたキーワードに割当て、当該キーワードの文字列と共にキーワードIDテーブル13に登録する。

ステップ305) 情報解析部12は、キーワードIDテーブル13に追加登録したキーワードについて、図13に示すテーブルのキーワードの出現頻度34を1にセットする。

【0057】ステップ306) 抽出したキーワードをキーワードIDテーブル13を参照してキーワードIDに変換し、図13で示したテーブルのキーワード出現頻度34を1増やす。

ステップ307) キーワードIDとその情報におけるキーワードの出現回数をそれぞれの情報プロフィール14に登録する。

【0058】適合スコア算出部16の処理は、図10に示す内容と同様であるが、その情報中のキーワードの出現回数を f_j ($j=1, 2, \dots, N$)、情報中のキーワードの重み V_j を以下のように計算する。

$$V_j = f_j / F_j$$

また、利用者プロフィール15の修正処理も図11に示した方法と同様の方法により行う。

【0059】

【実施例】以下、本発明の実施例を図面と共に説明する。

〔第1の実施例〕予め、センタ側装置10において、情報解析部12が情報蓄積部11からキーワードの出現回数、出現位置、文字数等を参照してキーワード『インターネット』を抽出し、当該キーワード『インターネット』に対応するキーワードID“01AC08”をキー

ワードIDテーブル13から取得し、当該キーワードの重み(1.000379)と共に情報プロフィール14に図5に示すように設定する(ステップ100)。

【0060】適合スコア算出部16は、図6に示す利用者プロフィール14から利用者毎に設定されているキーワードID『01AC08』とそのキーワードIDに対応するキーワードに関する重み(1.468824)を取得し、さらに、情報プロフィール14の情報から、抽出されたキーワード『インターネット』に対応する情報の適合スコアを算出する(ステップ110)。

【0061】提供情報作成部17は、適合スコア(S)に従って、利用者側装置20に送信する情報を選択し、利用者側装置20が表示可能な情報に編集し、編集された情報を利用者側装置20に送信する(ステップ120)。利用者側装置20では、センタ側装置10から取得したキーワード『インターネット』に対応する情報を表示し、当該表示情報に対する操作を行い、その操作状況を操作履歴テーブルに設定する。所定の周期または、任意のタイミングで操作履歴テーブルの情報をセンタ側装置10に転送する(ステップ130)。

【0062】センタ側装置10は操作履歴テーブルの情報を受け取ると、適合度算出部18において当該情報に基づいて適合度を算出し、利用者プロフィール修正部19において、算出された当該適合度に基づいて利用者プロフィール15を更新する。この例では、当該利用者の閲覧時間が所定の時間よりも長かった($a > 0$)ものとし、利用者プロフィール修正部19において、図6に示すキーワードID“01AC08”の重みを“1.789119”に更新する。

【0063】[第2の実施例] 予め、センタ側装置10において、情報解析部12が情報蓄積部11からキーワードの出現回数、出現位置、文字数等を参照してキーワード『デジタル』を抽出し、当該キーワード『デジタル』に対応するキーワードID“009001”をキーワードIDテーブル13から取得し、情報プロフィール14に設定する(ステップ100)。

【0064】適合スコア算出部16は、図14に示すキーワード『デジタル』の出現回数(16)とキーワード出現頻度記録部31から当該キーワード『デジタル』に対応する情報の重要度として重みを算出し、図6に示す利用者プロフィール15の重み(1.211203)を用いて適合スコア(S)を算出する(ステップ110)。

【0065】提供情報作成部17は、適合スコア(S)に従って、利用者側装置20に送信する情報を選択し、利用者側装置20が表示可能な情報に編集し、編集された情報を利用者側装置20に送信する(ステップ120)。利用者側装置20では、センタ側装置10から取得したキーワード『デジタル』に対応する情報を表示し、当該表示情報に対する操作を行い、その操作状況を

操作履歴テーブルに設定する。所定の周期または、任意のタイミングで操作履歴テーブルの情報をセンタ側装置10に転送する(ステップ130)。

【0066】センタ側装置10は操作履歴テーブルの情報を受け取ると、適合度算出部18では当該情報に基づいて適合度を算出し、利用者プロフィール修正部19において、算出された当該適合度に基づいて利用者プロフィール15を更新する。この例では、当該利用者の閲覧時間が所定の時間よりも長かった($a > 0$)ものとし、図6に示すキーワードID“009001”の重みを“1.339231”に更新する(ステップ140)。

【0067】なお、本発明は、上記の実施例に限定されことなく、特許請求の範囲内で種々変更・応用が可能である。

【0068】

【発明の効果】 上述のように、本発明の情報提供方法及びシステムによれば、キーワードをキーワードIDに変換し、利用者のプロフィールを提供するための情報のキーワードを全てキーワードIDとして取り扱うことによって、適合スコアを算出する際に、利用者プロフィール中のキーワードIDから蓄積されている全ての情報を検索し、その情報での当該キーワードの重みを参照する処理に要する時間が短縮され、その結果として、蓄積されたテキスト情報の中から利用者の興味に合致した情報を提供する従来の技術に比べて、より高速な処理が可能となり、情報の量が多い場合や、1つのセンタに対して利用者の数が多い場合に対応できる。

【0069】また、本発明によれば、1つの情報に対して情報中に含まれる多くの単語をキーワードとすることにより、より利用者の要望に合致し、利用者に対して漏れのない情報提供を行うことができる。さらに、センタ側装置において蓄積されている全情報、または、量的に十分な情報に含まれる当該キーワードの出現頻度を用いることにより、当該出現頻度から適合スコアを求めることにより、より出現頻度が高いキーワード(汎用的なキーワード)に対応する情報を利用者に提供することが可能である。

【図面の簡単な説明】

【図1】本発明の原理を説明するための図である。

【図2】本発明の原理構成図である。

【図3】本発明のセンタ側装置の構成図(その1)である。

【図4】本発明のキーワードIDテーブルのデータ構造の例である。

【図5】本発明の情報プロフィールのデータ構造の例である。

【図6】本発明の利用者プロフィールのデータ構造の例である。

【図7】本発明の利用者側装置の構成図である。

【図8】本発明の一連の動作の概要を示すフローチャー

トである。

【図9】本発明のセンタ側装置における第1の情報プロフィール作成処理のフローチャートである。

【図10】本発明の適合スコア算出処理のフローチャートである。

【図11】本発明の利用者プロフィールの修正処理のフローチャートである。

【図12】本発明のセンタ側装置の構成図(その2)である。

【図13】本発明のキーワードIDテーブルとキーワード出現頻度記録部を組み合わせたデータの構造である。

【図14】本発明の第2の情報プロフィールのデータ構造である。

【図15】本発明のセンタ側装置における第2の情報プロフィールの作成処理のフローチャートである。

【符号の説明】

- 10 センタ側装置
- 11 情報蓄積部、情報蓄積手段
- 12 情報解析部、情報プロフィール作成手段
- 13 キーワードIDテーブル
- 14 情報プロフィール
- 15 利用者プロフィール、利用者情報保持手段
- 16 適合スコア算出部、適合スコア算出手段

- 17 提供情報作成部、情報選択手段
- 18 適合度算出部、適合度算出手段
- 19 利用者プロフィール修正部、利用者情報更新手段
- 20 利用者側装置
- 21 情報表示部
- 22 操作部
- 23 履歴情報保持部、操作履歴保持手段
- 24 履歴情報転送手段
- 30 キーワードIDテーブル・キーワード出現頻度記録部
- 31 キーワード出現頻度記録部
- 32 キーワードID
- 33 キーワード
- 34 キーワード出現頻度
- 40 送信手段
- 131 キーワードID
- 132 キーワード
- 141 キーワードID
- 142 重み
- 143 キーワードID
- 144 出現回数
- 151 キーワードID
- 152 重み

【図1】

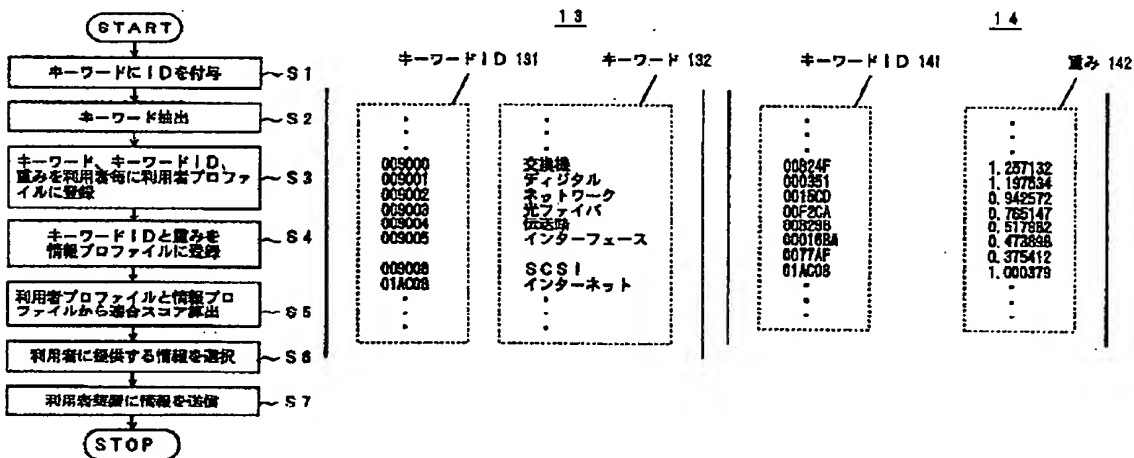
【図4】

【図5】

本発明の原理を説明するための図

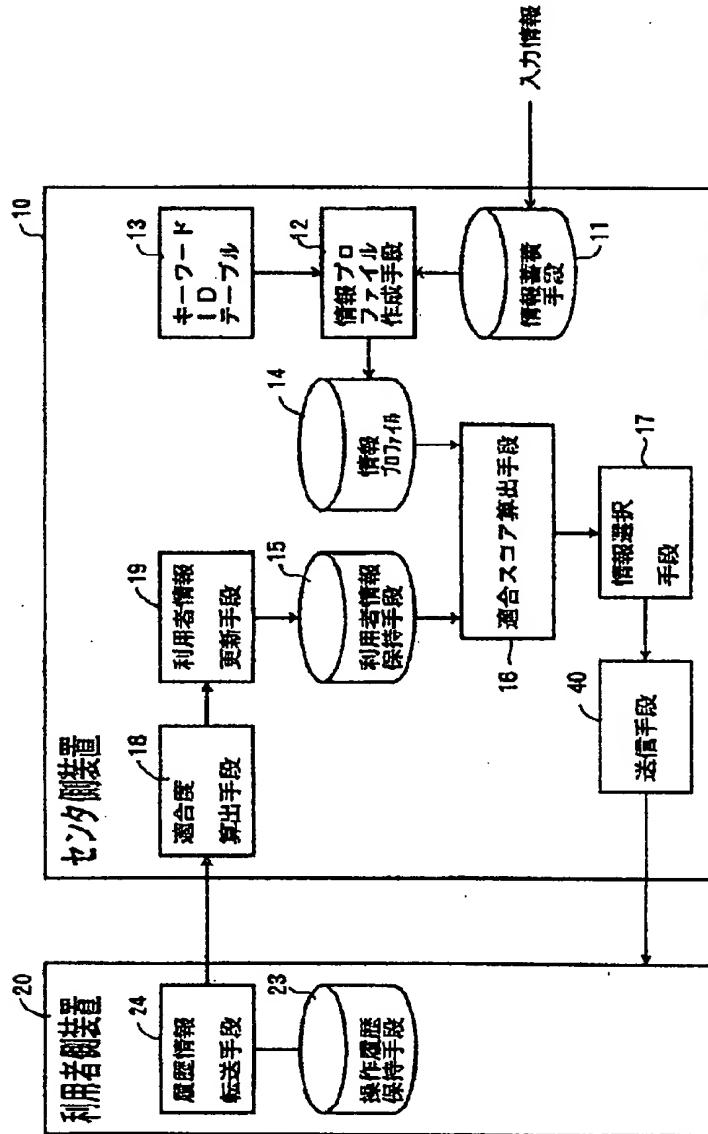
本発明のキーワードIDテーブルのデータ構造の例

本発明の情報プロフィールのデータ構造の例



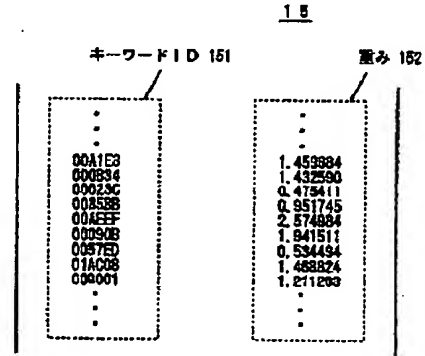
【図2】

本発明の原理構成図



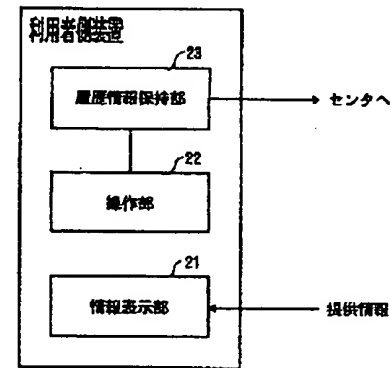
【図6】

本発明の利用者プロフィールのデータ構造の例



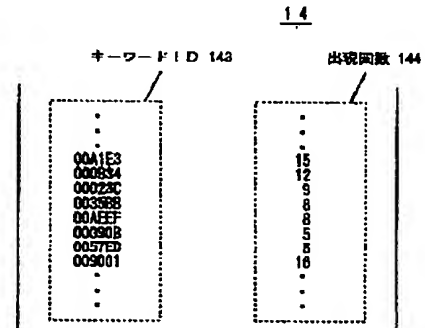
【図7】

本発明の利用者側装置の構成図



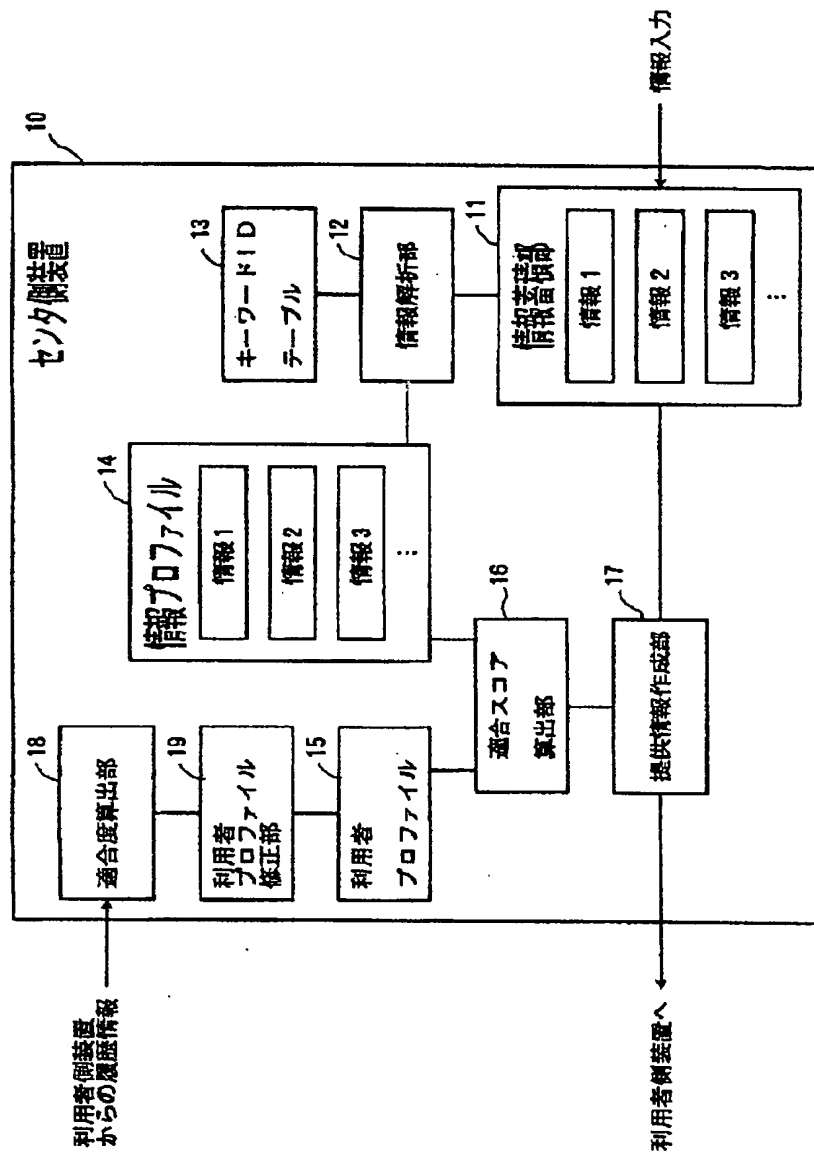
【図14】

本発明の第2の情報プロフィールのデータ構造の例



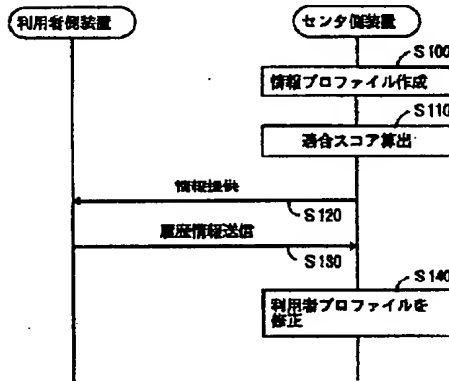
【図3】

本発明のセンタ側装置の構成図（その1）



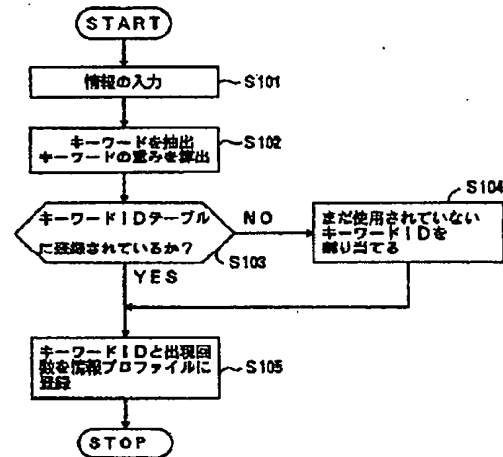
【図8】

本発明の一連の動作の概要を示すフローチャート



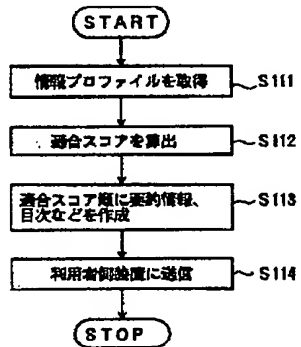
【図9】

本発明のセンタ側装置における第1の情報プロフィール作成処理のフローチャート



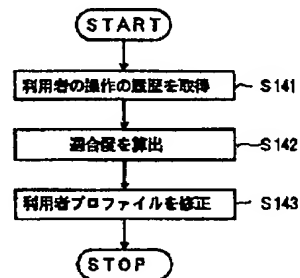
【図10】

本発明の適合スコア算出処理のフローチャート

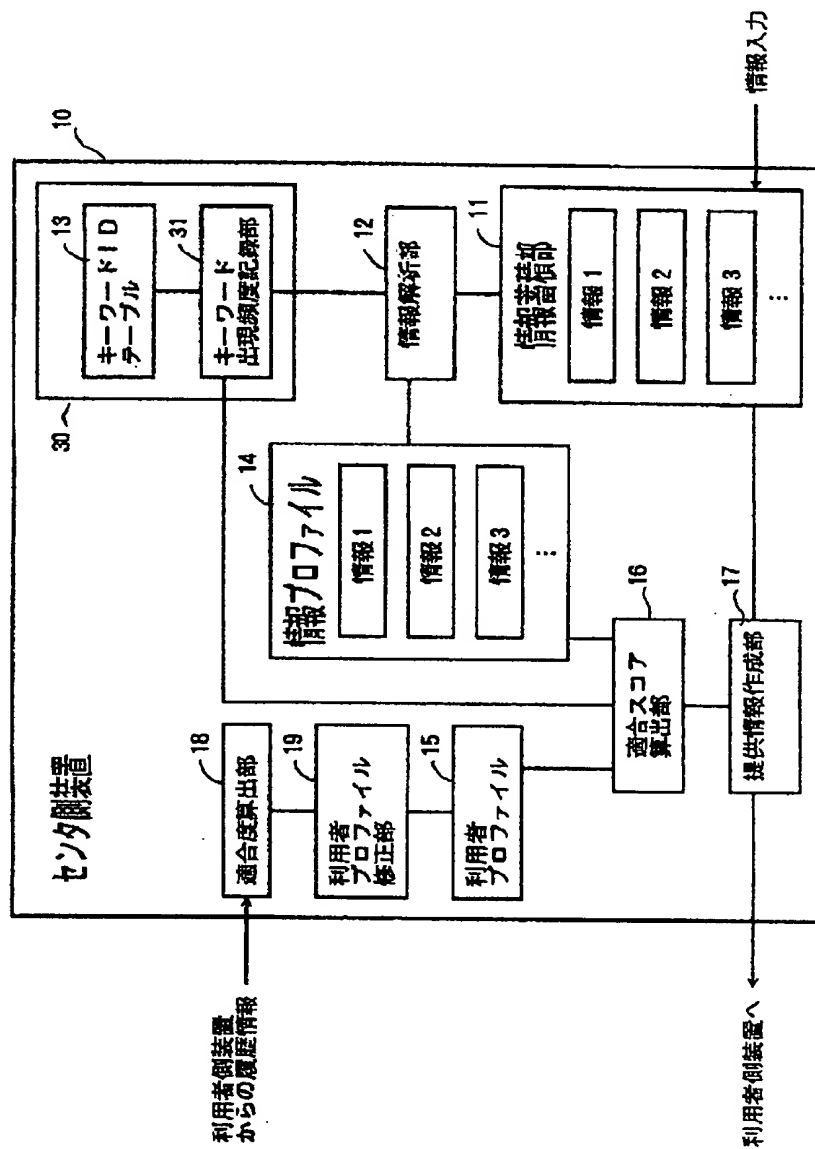


【図11】

本発明の利用者プロフィールの修正処理のフローチャート



本発明のセンタ側装置の構成図（その２）



【図13】

本発明のキーワードIDテーブルと
キーワード出現頻度テーブルを組み合わせたデータの構造の例

30

キーワードID 32	キーワード 33	キーワード出現頻度 34
...
009000	交換機	218
009001	デジタル	5878
009002	ネットワーク	1531
009003	光ファイバ	149
009004	伝送路	87
009005	インターフェース	212
009006	SCSI	28
...

【図15】

本発明のセンタ側装置における第2の情報プロフィールの
作成処理のフローチャート

